

Your Quick-Start Guide to Data-Driven Decision Making

Try Smartsheet for Free

In business, you constantly have to make decisions — from how much raw material to order to how to optimize retail traffic for changing weather. In days gone by, you might have consulted the person who had been around the longest for their best guess; for a more scientific approach, you might have also looked at sales records. Today, companies are finding that the best answers to these questions come from another source entirely: large amounts of data and computer-driven analysis that you rigorously leverage to make predictions. This is called *data-driven decision making* (DDDM).

To quickly get you up to speed with data-driven decision making, we've pulled together this action-oriented guide that explains the technology and includes expert tips as well as easy, step-by-step how-tos.

WHAT DOES DATA-BASED DECISION MAKING MEAN? WHAT IS DDDM?

Data-driven decision making (DDDM) is the practice of collecting data, analyzing it, and basing decisions on insights derived from the information. This process contrasts sharply with making decisions based on gut feeling, instinct, tradition, or theory. Data-driven decisions are more objective and can be easily evaluated according to their impact on metrics.

For example, a manufacturer that wants to increase production might time the assembly line, noting any delays. After introducing a solution to a bottleneck, the manager would time the process again to see if the change results in time savings. They might also try another solution and see which one actually speeds up production more.

The process becomes more sophisticated when you use data about the past to predict what will happen in the future.

This approach goes by many names, including big data, data analytics, business intelligence, diagnostic analytics, data analysis, data modeling, and online analytical processing.

Underlying DDDM is the belief that better data makes for better decisions. Without data, people run a much greater risk of being swayed by biases or acting on false assumptions. As the legendary engineer W. Edwards Deming put it, "Without data, you're just another person with an opinion."

Data-driven decision making is generating a lot of buzz in the business world as well as in the spheres of education, healthcare, and government policymaking.

But of course, the success of data-driven decision making depends on the quality of data collected and the methods used to sift through it. Data-driven decision making is heavily quantitative — historically, its use has been limited due to the need to perpetually collect statistics and crunch numbers. But the spread of moderately priced computing systems powerful enough to analyze large data sets — so-called big data — has made DDDM much more accessible.

THE BENEFITS OF DATA-DRIVEN DECISION MAKING

DDDM is not confined to big companies and government bodies with vast resources. Organizations of any size can use data-driven decision making to transform their processes.

With DDDM, organizations become more agile, detect new business opportunities sooner, and respond to market changes more quickly. With near-real-time data collection, managers can swiftly measure results and create a fast feedback loop. These abilities make data-driven companies exceptionally customer focused and more competitive. The same is true of organizations and service providers, such as public education agencies.

Data-driven decision making leads to greater transparency and accountability, and this approach can improve teamwork and staff engagement. DDDM policies make clear that whims or fads are not driving the organization, and morale improves because people see that objective data backs up management decisions.

In organizations that prioritize data-driven decision making, goals are concrete, and results are measured. Team members often feel a greater sense of control because they can see the goal posts clearly. The tenor of interactions may become more positive because discussions are fact-based, rather than driven by ego and personality.

Data-driven analysis can pay for itself through cost savings and higher revenues. Most organizations collect data, usually for record-keeping and compliance, but many don't do anything with this information. Often, they incur storage costs for keeping the data. So, why not sift through it and see if it tells any stories?

WHAT IS A DATA-DRIVEN APPROACH?

The term "data-driven" may seem redundant, as people who make decisions might already rely on data. But in the case of data-driven decision making, companies collect data methodically and analyze it rigorously, so that the information represents reality much more accurately.

Business managers typically rely on both qualitative and quantitative sources of information to make their decisions. The difference between this approach and data-driven decision making is how the data is collected and processed.

In the absence of a systematic process to gather and analyze data, what we may perceive as reliable data is actually a hodgepodge of anecdotal evidence, personal impressions, and selective information. This "data" consists partly of intuition and experience, and is not very scientific.

Of course, intuition does have its place in business. But, as an organization grows in scale and the dollar amounts at stake become larger, relying on gut instinct becomes very risky. As mentioned above, with data-driven decision making, the processes of data collection and analysis become methodical and rigorous, so that the information yields a far more accurate reflection of reality.

Common sense is important too, especially when bearing in mind that correlation does not equal causation. To wit, sage investor Warren Buffett once famously quipped about his fondness for junk food: "I checked the actuarial tables, and the lowest death rate is among six-year-olds. So, I decided to eat like a six-year-old."

When we talk about big dat

inately, give some people



“A big error we see businesses committing is going in 100 percent for data and throwing intuition by the wayside. While data is fantastic, it, by nature, can't think outside of the box,” notes Brandon Andersen, Chief Strategist at marketing analytics firm [Ceralytics](#). “That's up to humans. But, data is a great way to show whether or not what you're doing is working. And, it recognizes patterns that humans may be too busy to see. Data is a tool — an essential tool — but it's not the only tool for decision making,” he says.

anxiety, so the idea of data analytics can deter them.

[Business intelligence tools](#) alleviate some of this pain by automating data collection and hiding analytics under the hood. But simply acquiring these tools won't foster a culture of data-driven decision making any more than a paint-by-numbers kit will nurture artistic ability.

Creating a data-driven culture requires a long-term commitment to educating all members of an organization and championing this effort from the very top. Even then, expect the evolution toward this new way of doing business to take time.

This evolution toward a data-driven culture usually follows five stages:

- Data Denial: The organization starts with an active distrust of data and does not use it.
- Data Indifference: The company has no interest in whether data is collected or used.
- Data Aware: The business is collecting data and may use it for monitoring, but the organization does not base decisions on it.
- Data Informed: Managers use data selectively to aid decision making.
- Data Driven: Data plays a central role in as many decisions as possible across the organization.

(Note: Some strategists consider “data informed” to be the optimal state because it retains a bigger role for human judgment and balances the use of data with that of personal expertise and insights.)

As companies mature in the data-driven stage, they typically progress in their use of analytics from descriptive to diagnostic, predictive, and prescriptive.

Data-driven organizations share some key attributes:

- An emphasis on data collection
- An investment in tools and skills to make sense of data
- A commitment to making data widely accessible
- A willingness to consider data-driven ideas that arise at any level of the organization
- A dedication to ongoing improvement



Veteran Data Scientist Paco Nathan of [Derwen.ai](#) says that breaking down barriers is an important cultural shift for businesses as they transform into data-driven organizations. “The most common mistake is that organizations tend to react to complex problems by creating silos of people, processes, data, etc.,” he says.

“Silos are the enemy for data-driven business. They tend to create tech debt and other friction, which prevents firms from even beginning to work with data at scale. Getting effective data infrastructure in place, paying down the tech debt to do so, and removing the silos — those are the table stakes in this game,” Nathan continues.

In organizations, [change management](#) is becoming increasingly data driven, and analytics using complex data sets are playing an important role in effectively managing organizational change, [Harvard Business School professor Michael Tushman](#) and his colleagues have written.

This phenomenon is becoming more evident as big data facilitates continuous improvement processes in which companies implement incremental changes, monitor key metrics, and then make further adjustments based on the findings. Data-driven change management offers some benefits over traditional models. DDDM is less dependent on the expertise of a few key leaders and delivers higher-quality decision making based on facts, a greater capacity to scale changes, agility in modeling change scenarios, and the potential for rapid implementation.

WHO USES DATA-DRIVEN DECISION MAKING? EXAMPLES AND CASE STUDIES

DDDM can work for an organization of any size, from multinational behemoths to family businesses, as long as there's a commitment to the method's principles.

Big tech companies have pioneered and perfected DDDM. These firms possess a unique mix of analytical minds, technical expertise, and open culture that is conducive to data-driven decision making.

- **Facebook:** Facebook discovered early that democratizing access to data — that is, making it widely available — enabled the company to be much more agile and responsive to market changes and product development. In one example of the impact of this approach, Facebook looked at how many people started to use a feature through which they could ask a friend to take down a photo of them, but then abandoned that request when they realized they had to write a message to the friend explaining why. Facebook analysts found that if they auto-populated a sample message, the [number of users who completed the request rose to 60 percent from 20 percent](#). That data drove the decision to make the auto-populated message an official part of the tool.
- **Netflix:** A fascinating example of data-driven decision making comes from video-streaming service Netflix, a leader in original programming. (According to a 2018 Morgan Stanley study, 39 percent of U.S. consumers say Netflix has the best original programming — close to three times as many as their nearest rival, HBO.) What's Netflix's secret? The company used data analytics to build a highly detailed picture of consumers' tastes in videos. Then, they invested heavily in developing content that ticked off the viewer preference boxes. The result? *House of Cards*.
- **Google:** Then, there's Google's Project Oxygen, an initiative started in 2008 that mined data from performance appraisals, surveys, and other sources, with the aim of developing better managers. The project looked at the differences between the best and worst bosses in the company based on more than 100 variables. Google used the results to create training programs for managers and identified the eight most important behaviors of high-performing managers, starting with good coaching. The result was an improvement in management and outcomes, such as staff retention and satisfaction. (Several years later, Google updated the research, and in 2018, they refreshed the list of manager behaviors and added two more.)
- **Southwest Airlines:** More traditional businesses have also learned how to harness the power of analytics. For example, Southwest Airlines found that it could use analytics to save on jet fuel. The company also discovered that it could determine which airport gates were open to receive aircrafts so that customers would have to spend as little time waiting as possible.
- **Walmart:** Sometimes, the results of data-driven decision making are a little less glamorous, but nonetheless constitute home runs for a business. In 2004, for example, Walmart used product purchasing data from areas where hurricanes had struck in order to find out what people bought when they stocked up before a storm. The company wanted to use predictive analytics to determine how to supply stores ahead of future storms. They found that in addition to staples, like flashlights, stores saw heavy demand for unexpected items. Sales of strawberry Pop-Tarts jumped seven-fold as a tasty nonperishable that doesn't require cooking, and beer was the top-selling item. The retailer began sending trucks loaded with these items to stores in areas where hurricanes were forecast, and sales were brisk.
- **Amazon.com:** Under Jeff Bezos, Amazon uses a few simple guiding principles for management. To start, the e-commerce giant constantly tracks key performance indicators (KPIs) to measure its performance. The company also makes it easy for its people at every level (not just executives) to access data related to their roles. And when it comes to making decisions, Amazon's culture prioritizes data and evidence over seniority and personal influence.

Bezos has also realized that the company needs to use data wisely if that data is to supplement the decision-making process, rather than bog it down. As such, he advocates classifying a decision as non-reversible (Type 1) or reversible (Type 2) and devoting only as many resources to each type of decision as necessary. Bezos also suggests that trying to collect all the data needed to make a decision, especially if it's a Type 2, is unrealistic and expensive. Instead, he says, a preponderance of evidence (70 percent) is enough to get started. You can always correct course later. He accepts that even with DDDM, complete consensus on a decision is not always a given. So, he argues that managers should allow people to gamble on Type 2 decisions and learn from the aftermath.

- City of Boston: The city of Boston, Massachusetts has embraced data-driven management with a program called CityScore. The initiative consists of an online dashboard that shows how the government is performing relative to its goals in 24 key areas, such as responding to emergency calls and collecting garbage. Sensors automatically record much of the data, and city workers record information on mobile apps when they complete an activity. The system makes problems apparent and helps ensure the city allocates resources where they will have the most impact for citizens.

A SNAPSHOT OF DATA-DRIVEN DECISION MAKING IN BUSINESS

Despite these success stories, the [potential for data-driven decision making to make a bigger positive impact remains large](#) for many companies. About six in 10 companies responding to the survey said that their companies make a majority of decisions based on gut feel and experience, instead of data and information. Forty percent of best-in-class companies make decisions based on gut feel or experience, while, among laggards, the figure rises to 70 percent.

Less than half of the companies agree that information is “highly valued for decision making” or “treated as an asset in their organization” today. However, a full two thirds expect that it will be in the future. Only one third of enterprises use data to identify new business opportunities and predict market shifts. Again, though, most of the other two thirds say they want to do those things going forward. And on average, only about half of all available information is used for making decisions.

Andrew McAfee and Erik Brynjolfsson, professors at the MIT Sloan School of Management, conducted a survey and found that the [companies that were mostly data driven had six percent higher profits and four percent higher productivity](#) than average. While those might sound like small differences, the impact on results is substantial. Eighty-three percent of respondents to a survey by [The Economist said that using data had made their existing offerings more profitable](#).

KEY TERMS IN DATA-DRIVEN DECISION MAKING

Before we discuss how to get started with DDDM at your organization, you'll need to be familiar with some terminology.

- DataOps: A methodology for managing data and streamlining data analytics that combines data quality with integrity management, data security, and data engineering.
- Data Cleaning or Cleansing: The process of finding and either fixing or deleting data set records that are inaccurate, incomplete, or corrupt.

- **Data Governance:** Overseeing the integrity and security of data owned and used by an organization, as well as managing that data's availability and type of use.
- **Data Warehouse:** An integrated, centralized store of all the data an organization collects and uses for keeping records, reporting, and performing analytics.
- **Data Mart:** A subunit of the data warehouse that deals with a single department at an organization.
- **Descriptive Analytics:** The preliminary stage of data processing in which one extracts historical insights from data and prepares it for more advanced forms of analysis.
- **Diagnostic Analytics:** The branch of data analytics that focuses on determining the causes of phenomena.
- **Predictive Analytics:** The branch of data analytics that focuses on extracting patterns from historical data with the aim of predicting future events.
- **Prescriptive Analytics:** The branch of data analytics that focuses on using data to determine the most appropriate course of action when there's a decision to make.
- **Four Vs of Big Data:** The four Vs are the four critical characteristics of big data. *Volume* refers to the amount of data. *Velocity* is the speed at which data is collected and/or analyzed. *Variety* is the number of different sources from which data is accumulated. And, *Veracity* is a measure of how trustworthy the data is.
- **Database:** An organized collection of data records that can be accessed through a computer system.
- **Relational Database:** This is a type of database that can recognize pieces of data in relation to other pieces of data in the same database. This type is of limited utility for big data sets, as its need for structured data means that performance slows down as the size of the data increases.
- **NoSQL:** A big data-friendly database that, thanks to its unstructured nature, can maintain high performance when dealing with very large data sets.
- **Hadoop:** This is an Apache software ecosystem that enables parallel computing on a very large scale by allowing data to be spread across a large number of servers. Hadoop can radically reduce data-processing times.
- **CAP Theorem:** This refers to the idea that a distributed database can only have two of the following three qualities: consistency, availability, and partition tolerance. There are a couple of alternative philosophies for data system design that illustrate this principle: ACID prioritizes consistency, while BASE prioritizes availability.
- **Data Point:** A measurement derived from a single member of the units being observed (in other words, a fact).
- **Data Visualization:** A data presentation technique that uses visual tools to help people understand data in context.
- **Data Gap Analysis:** A technique that uses analyses of existing data to determine where an organization is failing to collect or analyze other potentially useful data.
- **Key Performance Indicator (KPI):** A defined measurement, collected regularly, that offers a snapshot of organizational performance in a specific area.
- **Data Model:** A data model defines how various points of data within a database are connected,

processed, and stored in relation to one another.

- **Metadata:** This refers to data about data. Metadata records information about other data.
- **Ad Hoc Analysis:** An analytics process built to answer a specific business question.
- **Application Programming Interface (API):** A software mechanism that allows applications to access features of other services, such as operating systems and other applications.
- **Artificial Intelligence:** The constructed ability of machines to simulate human intelligence.
- **Machine Learning:** A branch of artificial intelligence in which computers have the capacity to learn by figuring out patterns in data.
- **Contextual Data:** Data that allows people, items, or other entities to be considered in context (as part of a bigger picture).
- **Extract, Transform, Load (ETL):** A combined set of database functions that extract data from one database and put it into another data repository.
- **Hierarchy:** The organization of data points into a systematic structure.
- **Slice and Dice:** The process of breaking up data into its constituent parts to facilitate the extraction of information.
- **Zero-Latency Enterprise:** This refers to data architecture with the ability to exchange information across technical (OSs, DBMSs, programming languages) and organizational boundaries in near-real time, with negligible delays in transmission. Latency is the delay between an instruction for the reception of a transfer and the sending of a transfer.
- **Insight Platforms-as-a-Service (IPaaS) –** This refers to the cloud-hosted big data services that are intended to make data analytics more widely accessible. They provide the tools needed to parse insights from data.
- **Translytics:** A category of databases that can maintain high performance when dealing with both transactions and analytics.

A STEP-BY-STEP GUIDE TO START DATA-DRIVEN DECISION MAKING

Now that you've learned where and why organizations are using DDDM, you may want to start implementing this approach at your own enterprise. Large amounts of data can quickly become overwhelming, so follow these steps to get going without headaches.

7 Steps of Data Driven Decision Making

1

Step 1:
Define objective



2

Step 2:
Establish hypothesis



3

Step 3:
Identify data need



4

Step 4:
Build data process



5

**Step 5:
Collect
data**



6

**Step 6:
Analyze
data**



7

**Step 7:
Make decision**

Use this downloadable checklist to help guide your process. The sheet walks you through important steps, such as forming your team and communicating your decisions.

DATA-DRIVEN DECISION MAKING CHECKLIST

Stage One – Define Objectives and Get Going

- 1. Identify the roles and people you need for this project, especially any specialized talent.
- 2. Make sure everyone understands the basics of DDDM and buys into the value of data-driven decision making.
- 3. Define the goals of your project. Articulate them in specific, measurable terms such as, "We seek a 20 percent reduction in customer returns over six months."

Stage Two – Frame the Assignment

- 1. Identify the area of your business where your effort will have the biggest impact.
- 2. Articulate the key question in the form of a hypothesis, an if-then statement that can be proven or disproven.

Stage Three – Data Process

- 1. Identify the data you need
 - a. Qualitative vs Quantitative
 - b. How many sources of data will you need?
 - c. Will it be reliable and comparable?
 - d. Are the benefits worth the costs?
 - e. How long will it take to get the data?
 - f. How much of the data are you already collecting?

Stage Four – Filling Data Holes

- 1. Where can you obtain data that you do not have?
- 2. Set a budget for this.
- 3. Get necessary approvals
- 4. Identify any service providers needed and engage them.

Stage Five – Data Collection

- 1. Assign people to data collection and management roles.
- 2. Define the process and protocols for data collection.
- 3. Clean and preliminarily process data.

Stage Six – Data Analysis

- 1. Using analytics, test your hypothesis.
- 2. Model different scenarios based on changing variables.
- 3. Identify potential change actions.

Stage Seven – Present Findings and Decisions

- 1. Create a presentation of your findings using visualizations and graphics where possible.
- 2. Articulate concrete benefits for proposed changes.
- 3. Make decisions.
- 4. Form an implementation plan.
- 5. Communicate with key stakeholders.

Download Data-Driven Decision Making Checklist

[Excel](#) | [PDF](#)

Steps to Implement Data-Driven Decision Making at Your Organization

Following are the seven steps to implement data-driven decision making at your organization:

- Step 1: Start by Defining the Objective and by Fostering a Culture: First, figure out what goals your business hopes to achieve through DDDM. Identify your business goals, and then build a broad

strategy around them. At this stage, it's fine if the strategy isn't completely defined. But, you do need to know how you think big data will help achieve your targeted outcomes.

For example, if you own a big chain of grocery stores, you can examine customers' fruit purchases with the goals of improving the supply chain, the placement of products in the store, the pricing, and even the packaging of store-brand offerings. Articulating these goals will shape your data collection and analytics strategies from the outset.

Also at this stage, you'll want to lay the groundwork for success by communicating to all members of your organization your commitment to data-driven decision making. This shared value and culture is just as important as having the infrastructure and tools in place. By committing to using data, you're acknowledging its effectiveness — and you have to do so in a way that breeds confidence in everyone else.

Determine whether your goals require using real-world data for analytics or experimental data for simulations or modeling. *Real-world data* is what you're collecting and analyzing in the grocery store example, where you're trying to figure out what is happening in an existing business. *Experimental data*, on the other hand, is collected from a controlled environment with the aim of figuring out what business practices work best.

- **Step 2: Focus on a Specific Business Area and Define Questions:** Determine which area of your business needs the most attention. Typically, the answer is customers, finances, or operations (or some mixture of the three). Which do you suspect will bring the most benefits in terms of business growth? Prioritize the area that will have the biggest impact.

Your next step is to identify the specific business questions you want to answer. Which of these are related to the business goals we talked about in step one? Formulate this thought in the form of a hypothesis, a proposed if-then statement that can be proven or disproven and that you can use as a starting point for investigation. For example, "If we add a self-service portal to the website, we will decrease the number of after-hours emergency trouble calls."

Now, you'll also know the specific data sources you need to focus on and which data you can hope to collect from them.

When you identify in advance the specific data you need to collect and how you're going to collect it, you spend much less money, time, and stress than if you were simply to collect everything and try to figure out what to do with it later.

- **Step 3: Identify the Data You Need:** The data you examine needs to be relevant to your business question. To build on our grocery store example, customer payment receipt data would be useful; ambient store temperatures, not so much.

There are two broad categories of data, and since each informs decision making, you'll want to collect both:

- **Qualitative Data:** This kind of data is non-numeric and subjective. It is observed, rather than measured. For example, having a store worker ask people a few questions about their shopping experience will generate qualitative data. The shoppers may say things like, "The meat section is too far from the wine section" or "I have to make my trips super short because my toddler won't sit still in this shopping cart seat."

- **Quantitative Data:** This kind of data is numerical and objective and is measured, not observed. Quantitative data is what people think about when they think of big data: rows upon rows of data concerning who purchased what, how much they spent on it, when they bought it, how often they buy it, and so on.

Not all data will be suitable for use. The data must measure something meaningful, and must also be accurate, formatted correctly, and not contain duplicates. For these reasons, data collection methods often undergo a pilot test before they're rolled out in force. If you're working from existing data, you may need to "cleanse" the data before conducting analytics, which involves verifying information, fixing errors, and removing redundancies.

To get a sense of your data's quality, here are some key questions to answer:

- Who collected the data, and is it reliable? Reliability requires that it measure what it claims to, that all data was collected using the same parameters, and that it does not contain false responses (such as a fake email address).
- How was the data sampled? Is the data representative of reality? Or, were only certain customers asked to respond?
- Does the data include outliers (exceptionally high or low measurements)? How are these affecting the overall data distribution?
- Are relationships that you see as causal truly causal? Or, are confounding factors and intermediate variables giving an inaccurate impression of causality?
- What are the assumptions underlying your data, and do they hold up to scrutiny? For example, we might assume that people buying groceries are buying them mainly for themselves, but we don't know how true that is, since shoppers are often running errands for other people.
- And, most crucially, why have you decided to analyze data one way and not another?

Even if you end up ruling out some data, you're likely to find that you have several internal data sources; one survey found that the [median number of data sources for companies is five](#). Companies on average have three external data sources. What's more, half of the people who work in data-driven businesses say the number of data sources they use is growing. Having multiple sources of data means that you can choose the best option for collecting data, based on whichever is fastest, cheapest, or least onerous.

- **Step 4: Determine How You Will Get the Data You Do Not Have:** You may find that you already collect some of the data you need. If this is the case, count yourself lucky: You will save on data collection costs, and you can begin analysis sooner.

If you do not already have the data, you have to figure out how to get it. Think of the four Vs of big data: How can you optimize volume, velocity, variety, and veracity? Is it by collecting the data yourself or by getting it from an external source?

Remember that holes in the data are more likely to crop up when different departments use different systems to collect, manage, and report that data. Some systems, which cross-analyze data that comes from different sources, can help solve this problem.

Having a variety of data sources is problematic if the data does not use common variables, which

enable you to integrate those sources. If the data does not use common variables, Northeastern University instructor Joel Schwartz encourages you to consider whether the data could be useful in other future projects. If so, it may be worth the investment of time and resources to solve these data disparities now.

Like any investment decision, big data demands that you have a clear case that weighs the costs against the expected benefits. Data is cheap, and it's getting cheaper. But, it's not free, especially if you're collecting it yourself, and storing it incurs costs well into the future.

That's why it's important that you focus only on the data that stands to deliver meaningful value. And, that process requires some homework. If obtaining and analyzing the data is going to cost more than the potential benefit, it's a bad business decision.

Here's a helpful tip: There are ways to reduce data costs, such as by looking at alternate sources or alternate data collection methods. However, if you've reached this stage without considering how you're going to benefit, even cheap data isn't going to turn a bad investment into a good one.

- Step 5: Collect the Data: When we talk about big data, we're usually talking about the automated data collection of thousands, millions, or even more cases. But, data analytics can also work with smaller quantities of data that are collected by people. In either scenario, there are three things to be done:
 1. Assign People to Data Collection and Data Management Roles: These people are the first — and sometimes the only — check on the veracity of the data.
 2. Define the Processes and Protocols of Data Collection: This is a critical step, because the methods of data collection have a direct bearing on the four Vs of the data. Identify data sources and spell out the precise methods of data collection — that is, whether the data is being recorded by a machine or by a human. You should also create a data dictionary to catalog and define each of the recorded variables.
 3. Clean and Do a Preliminary Processing of the Data: Cleaning data is arduous but vital work to ensure integrity and usability. This process looks for duplicated, incorrect, or corrupted data. The analyst may also run preliminary descriptive analyses, such as data distributions, which will inform more advanced analytics.

Data is useful when you have a variety of it, and that means gathering data from different sources. Data from a single source, or even data from similar sources, is unidimensional, offering only a limited number of ways to think about the data.

You can also find or buy pre-cleaned data sets. These are of variable utility, depending primarily on the data's relevance to your goals.

"The most common error that we see in implementing data-driven decisions is neglecting to validate the data before drawing conclusions. It's so easy, when the data is available and ready, to assume it's correct and immediately actionable, but this is often not the case," explains Sam Underwood, Vice President of Business Strategy at data analytics firm [Futurety](#).

"Especially in organizations that have not previously relied heavily on data to drive action, this organizational data may not have undergone a quality control or governance process, and it is likely to be incorrect in many cases. We estimate that when we inherit our clients' data, around 50 percent of it has major accuracy issues. We often build in a two-to-four-month data-cleaning phase before committing to any deliverables that use the data," Underwood says.

- Step 6: Analyze the Data: You are finally getting close to realizing insights from your data. There are a number of platforms available for big data analysis, and people with a strong knowledge of Excel can use some of them. But, more advanced and potentially more valuable analytics call for a trained specialist. Add that to the costs of dabbling in big data.

Here's a helpful tip: Even if you do not have the skill to do this analysis yourself, it's still a good idea to know what type of analytics you need (such as diagnostic or predictive), because analyzing different data types calls for different skill sets.

Dealing with quantitative data is relatively straightforward, but qualitative data, like images, video, speech, and text, calls for different conceptual knowledge and technical skills.

- Step 7: Present the Findings and Make Decisions: Presenting the findings is a surprisingly difficult part of data analytics. After someone has crunched the numbers to come up with insights, you must deliver these findings to the right people at the right time in the right way.

Luckily, there's no shortage of interesting ways to present data. Even complex ideas can be easily communicated through good writing, clear presentation, and helpful graphics or visualizations.

The aim is to connect the insights you gained to actions that will benefit the business. Remember to clearly portray both the reward and the risk of big data-driven decisions. As David Ritter writes for [Harvard Business Review](#), *if the risk of acting on something and being wrong is too high*, acting on even firm insights can be a mistake.



Based on your findings, formulate action recommendations and an implementation strategy. A common pitfall is “not setting clear expectations around the decision-making process,” says James Horey, CEO of [Reviewbox](#), an application that helps vendors manage product information on e-commerce sites. He continues, “For example, what do they expect team members to bring to the table? How will decisions actually be made? When will there be ‘enough’ data to make the decision? The best way to overcome this is to come up with, document, and communicate this process to all the relevant team members.”

He adds, “As the data becomes easier to collect, the temptation increases to continually pass the buck and conduct more research, which leads to decision paralysis. Effective organizations understand that most decisions need to be made in a timely manner, and they, therefore, often employ an 80 percent

rule: We'll make the decision once we're confident that we have about 80 percent of the available information."

A helpful tip: You can help data analysts make their findings easier to understand by bringing in communication specialists, such as graphic designers and writers, at the presentation stage. They can assist with distilling insights and conveying them in impactful ways.

Formulate an implementation strategy for your decisions and communicate this with all stakeholders.

MANAGING TEAM MEMBERS WHO ARE DATA AVERSE

Despite your enthusiasm for data, you will probably find people in your organization who remain reluctant to embrace it — and not always for the wrong reasons. Why are people averse to big data, and what can executives and managers do to overcome this resistance?



"Frequently, upper-level management can have a negative effect on the right use of big data," says Carl Hillier, Head of Product Marketing for [Ephesoft](#) software, which uses machine learning to capture unstructured data. He adds, "These managers may resist making decisions based on data, which can undermine your data-related investments. That's why it's important to encourage team members and peers to use data to test intuition and solidify ideas."

Thomas C. Redman, President of Data Quality Solutions, says that [people aren't apathetic toward big data as much as they are afraid of it](#), as analytics and data-driven predictions move from specialist areas into the mainstream. Data now feels pervasive and is transforming work and power structures, causing many people to fear for their jobs. This reaction is perhaps not surprising, considering that people like venture capitalist Mark Kvamme have predicted that 20 million jobs will disappear from the United States as a result of artificial intelligence.

Others fear that big data is a false hope because the expertise to realize its potential is not sufficiently available at many companies. Some naysayers, remembering times when data led them astray, fear catastrophic mistakes.

Moreover, there are a number of less laudable concerns. Big data generally accompanies a shift toward

greater accountability for results, and that may shine an uncomfortable spotlight on people who try to get by with the bare minimum of work.

Managers and business leaders can allay many of those more reasonable fears from the outset, so that DDDM does not provoke resistance. Here are some essential tips:

- **Don't Frame Data Knowledge as an Arcane Skill Set:** If you do, you foster a belief that data is strictly the domain of gifted mathematicians and quants. Instead, invest in teaching people about analytics, and encourage them to use data to answer simple questions. (For example, "how often do team meetings start on time?") The more they learn about analytics, the less likely they will be to fear data. Lead by example and expand your own skills. Read a book on the subject, such as *A Practitioner's Guide to Business Analytics* by Randy Bartlett or *Keeping Up with the Quants* by Thomas Davenport. Enroll in one of the classes listed at the end of this article. Practice collecting some simple data, learning how to graph it, and performing some basic statistical calculations, such as figuring out the mean, median, mode, variance, standard deviation, correlation, and probability.
- **Don't Overpromise:** With both the C-suite and front-line staff, you want to set realistic expectations about what you can accomplish with data-driven decision making and how long it will take to get there. Present data as a tool, albeit a powerful one, not an earth-shattering paradigm shift. Manage expectations.
- **Show Data in Action:** Look for opportunities to show how data not only helps the business make more money, but also helps individuals do their jobs better by saving time and resources or decreasing stress. (For example, data shows how much traffic a store has gotten on past Super Bowl Sundays, so you can schedule enough people to provide good service, but not so many that staff are idle.) Publicize successful use cases. Don't forget to nurture champions for data — those individuals who act as leaders on your journey toward DDDM.

COMMON PITFALLS IN DATA-DRIVEN DECISION MAKING AND HOW TO AVOID THEM

As using data has helped businesses make better decisions, enthusiasm for DDDM has grown. Your information is an asset and a resource, but it's crucial to remember that data is only as valuable as the benefits you derive from it.

The key to making data worth its costs is identifying how you hope to use data and then figuring out what data you need. Working backward by gathering data and then trying to find a use for it is wasteful at best and useless at worst. Unfortunately, one of the leading complaints is that decisions were made without data because the right data was not available.

Another pitfall involves the inherent complexities of big data. Collecting massive amount of information, integrating it, extracting insights, and turning those insights into action-oriented recommendations in real time present challenges. So, expect to put a lot of work into designing your process, and realize that errors can arise at any stage. Failure to spot them can result in problems or poor decisions, so even with analytics, you can't work on autopilot.

Human nature also confounds our attempts to make decisions based on data, as we have a maddening tendency to favor our guts over the numbers. While this can sometimes lead to home runs, the tendency subverts the intention of DDDM.

Human cognitive biases can muddy the waters further. We may cherry-pick numbers, falling victim to confirmation bias. Or, the highest paid person's opinion (the delightfully named HiPPO) may sway us, or we



Derek Wilson, CEO of predictive analytics firm [CDO Advisors](#), says that companies new to DDDM can run into problems by jumping into the deep end. “Start with easy-to-understand problems and solutions. Team members need to be able to understand the data and decisions,” he urges. A common mistake, he says, is “businesses that start with too many process changes at once. Implement data-driven decisions one by one and get quick wins to build confidence in the team.”

may subscribe to the same assumptions about the data as everyone else (groupthink).

There’s a litany of other phenomena that skews our thinking, from loss aversion and the *sunk cost fallacy* (making a decision based on not wanting to waste resources that have already been spent and cannot be recovered) to the *Texas sharpshooter fallacy* (when differences in data are ignored, but similarities are stressed, leading to false conclusions) and the *IKEA effect* (a cognitive bias in which we place an irrationally high value on things we had a hand in creating).

As a final caution, critics say that over-reliance on data may go too far in removing human judgment from decisions or may dehumanize the subjects of the data, turning staff and customers into “just a number.” Many people also have privacy concerns around large-scale data collection and storage.

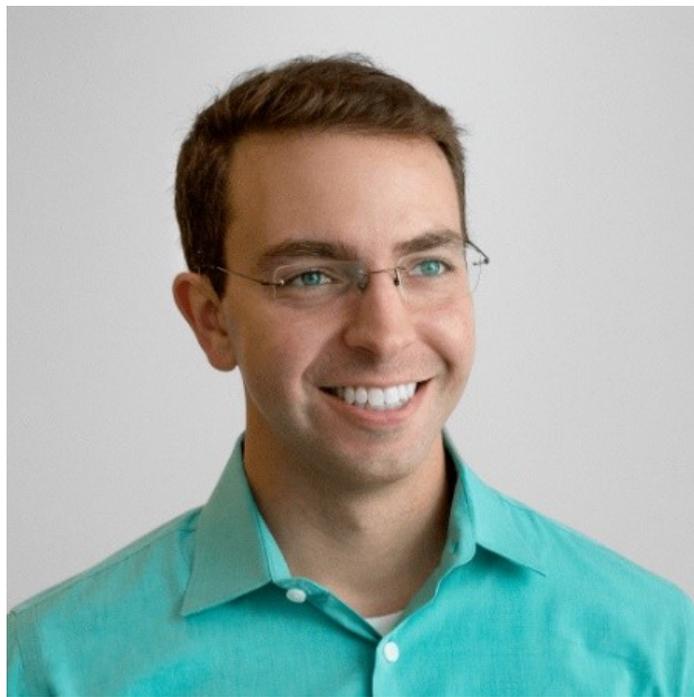
BUILDING A DATA-DRIVEN BUSINESS CULTURE

Organizations that have successfully built a culture founded on data-driven decision making say the most important thing to remember is that your company will not transform overnight. As Klipfolio’s Marketing Director Jonathan Milne says, [you have to “be patient, take your time, and start small.”](#)

To create the conditions for a data-driven culture, you must do the following: collect high-quality data within a useful timeframe; make sure the costs of your information architecture are reasonable; and be sure the data is easily accessible. Some experts urge companies to begin collecting data as early as possible, so they have benchmarks, while others caution against a scattershot approach.

Moreover, you need to know how to present, visualize, and talk about data. Create an internal data dictionary that outlines the content, format, and structure of your databases. Make sure that managers have at least basic statistics training.

Your management style must also change as your culture evolves. Data scientists Jeff Bladt and Bob Filbin say that in organizations making the transition to being data driven, there are always people who attempt to subvert the change and those who embrace it. [People who adopt the data-driven approach](#) do so to improve either their perceived or their actual performance, Bladt and Filbin say. People who have low



Tom O'Neill, CTO and Founder of business intelligence platform [Periscope Data](#), has worked with more than 1,000 teams of data scientists and offers some advice based on the most successful teams:

- Train everyone at your company to use and interpret data accurately. Recognize that no one is born data literate. All employees should know how to determine if they're using metrics, goals, and conclusions correctly.
- Analysts must work tightly with their business counterparts and drive change together. Intentionally include different levels of experience and perspectives throughout the process to reduce the biases that homogenous groups exhibit. Using this mix of people will minimize blind spots and maximize potential solutions.
- Encourage curiosity in your culture. Everyone should be empowered (and expected) to ask a lot of questions, and there need to be resources available to provide answers. You should regularly dissect KPIs into their components to provide new perspectives on what the non-curious might see as a simple sum or projection.
- Use data to make testable, tactical predictions. Then, take actions based on those predictions and feed the results back into your process to improve future decisions. The data-driven decision engine looks a lot like the scientific method. Historical analysis should be focused on learning why something happened, not just reporting what happened.

organizational prestige but are high performers will welcome the use of data, believing that it will make their contributions more evident. With encouragement, these individuals can turn into data champions.

Conversely, highly regarded team members who are actually poor performers remain distrustful of data, which they fear will reveal their shortcomings. Many of these people are unlikely to change their views, and, over the long run, they will probably need to leave the company for it to become truly data driven.

Data-driven decision making needs to be embraced at the upper levels if it's going to take root in the organization. Given that the successful implementation of DDDM requires particular qualities, such as a collaborative style and an openness to solutions that may come from anywhere in the ranks, the process itself demands practitioners who model from the top down and across departments.

Encourage consistent metric tracking (such as KPIs) throughout your organization as a way to get people to

focus on the relevance of data to business objectives. Building a [dashboard that gives you a fast snapshot of those KPIs](#) can aid this effort.

In addition, you may be wondering where in your organization responsibility for data-driven business intelligence should reside. (This responsibility is also referred to as [data governance](#).) Cross-departmental data “competency centers” are much more likely to be responsible for data in best-in-class companies (52 percent) than they are in average companies (40 percent) and stragglers (28 percent), according to a [survey regarding corporate data governance](#). When asked, “Who manages and governs data for decision making?” 60 percent of respondents said it was the finance department, while 41 percent said it was the IT department, and, again, 41 percent said it was a competency center.

Strong data governance assists DDDM. You can support that governance by fostering agile IT infrastructure, cross-departmental data management, a collaborative decision culture, and the use of KPIs throughout your organization.

THE ORIGIN OF DATA-DRIVEN MANAGEMENT AND THE TECHNOLOGIES BEHIND IT

DDDM became possible with the advent of databases that were capable of collecting and storing large amounts of information and the arrival of systems that could crunch this data. The emergence of networked computers and the internet has also played a major supporting role in the growth of big data.

The term big data began circulating around 2005, but the concept of using data to track and control operations has been around since the first entrepreneurial activities of ancient Egypt. Systems more akin to what we have today arose in the late 1930s, when IBM developed a punch-card reading machine to keep records for the new Social Security System. In the 1960s, electronic records storage began.

In the early 2000s, big data was in its infancy, and analytics consisted primarily of ad hoc data manipulation summarized in reports and KPIs. As the volume of data grew, the quantity of information became overwhelming, and data analysts bumped up against limitations. They found themselves unable to drill down to see which branches were underperforming or to compare performance to historical trends.

The introduction of Hadoop in 2005 marked a major milestone in the development of big data. Hadoop is an open-source operating system that performs the distributed parallel processing of vast quantities of data. Hadoop was invented to empower the analysis of large data sets; its ultimate goal was to be able to index the World Wide Web.

Hadoop became the backbone of big data operations and gave rise to sophisticated analytics operations that included interactive dashboards and data warehouses. Inspired by [Agile software techniques](#) and [DevOps](#), *DataOps* began to spread around 2017 (a methodology to improve the quality and speed of data analytics). Simultaneous with these improvements, the volume of information has continued to explode. As of mid-2018, an estimated 4.4 trillion gigabytes of digital information has been stored and made accessible to web users, with 90 percent of that having been created in the preceding two years.

Analytics has changed a lot in a short time. Today, the process is much more user friendly. Data collection is largely automated, and, while setting up that automation is still a specialist’s job, it is far easier than it used to be. Dashboards and data visualizations make data much easier to grasp, and generating a report is much less laborious and complex than it used to be.

Cloud solutions are making data analytics more accessible to organizations without traditional data expertise or experience.

To understand how we got here, you need to trace the developments of databases and data warehouses since the turn of the millennium. Early on, data would be stored in structures called *operational databases*, or *online transaction processing (OLTP) databases*, which were designed for *transactional systems*, such as the withdrawal of money from a bank account.

But, these transactional databases didn't work well for analytics, so *data warehouses* arose. A data warehouse, or *online analytical processing (OLAP) database*, is a database that is layered on top of an OLTP. This additional layer allows analytics to run efficiently. To move data from OLTPs to OLAPs, a process called *extract, transform, load (ETL)* comes into play. It extracts data from a source, transforms it into an analysis-suited format, and loads it into the data warehouse.

But, ETL was cumbersome, and the problem grew as the 4 Vs of data began expanding. Older database structures, called *SQL* or *relational databases*, couldn't handle the increased demands. As data turned into big data, we needed storage and analytical technologies that could keep up with the growth.

The answer came in the form of *NoSQL* databases and Apache Hadoop open-source software utilities for the distributed storing and processing of data. SQL databases only work with predictable, structured data, and to crunch more data, you need more computing power. NoSQL databases work with unstructured data, and you can handle more data by distributing the load over more ordinary servers. Also, during transaction processing, SQL databases prioritize properties summarized by the acronym ACID, whereas NoSQL databases work on the BASE model. While the intricacies are complicated, the main takeaway is that ACID is exceedingly structured and rigid. In addition, this model does not work well with big data, distributed computing systems, and unstructured data. The BASE model provides the needed flexibility.

With NoSQL came Hadoop, which, in simple terms, is a processing framework that allows for parallel computing. In other words, it has the ability to solve massive computations in parallel over a network of computers. Hadoop, in conjunction with NoSQL, is the superhero of big data analytics: This combination is cheaper than data warehousing, it's not limited to structured data, and it's suited for both storage and computation.

As it played out, however, an even better solution turned out to be a combination of Hadoop and data warehousing: Hadoop provided the capacity to store incredibly large quantities of data, and for analytics, ETL moved data from Hadoop into warehouses.

As data began to amass in NoSQL databases, the ETL process for analyzing it became less and less efficient, so analytics for NoSQL were developed. When these ran against constraints, the workaround became a combination of SQL and NoSQL for storage and analytics. A more recent development is *translytics*, database solutions that handle both operations and analytics with ease.

We can now do more with our data. As big data has matured, we are moving along the continuum from descriptive to diagnostic and predictive analytics, with prescriptive analytics increasingly within reach. The most familiar big data application is predictive analytics: the use of historical data to detect patterns and forecast data. Predictive analytics' combination of increased storage and increased computational power allows us to use machine learning.

Tech analyst Rob Enderle [says](#), "Firms that can change their practices to make decisions data driven will, as these practices shift, quickly surpass those organizations that do not."

If your organization has been slow to start with DDDM, perhaps you're concerned that your company will be left behind. But, there is a way to jumpstart your effort — by using cloud-based analytics. These platforms, called IPaaS (insight platforms as a service), have recently proliferated. These applications are designed to handle the integration of data and the generation of intelligent insights. Outsourced cloud solutions are not as customized as something designed in house, but they are less costly. And, if you are

struggling to get going, these tools are efficient. *Embedded analytics*, so named because the analytics functionalities are built into business applications, are another option. Both of these enable you to harness data at a low starting cost and with minimal data expertise.

A final technological development to be aware of is the attempt to make analytics work in real time, minimizing the lag between the generation of data and the generation of insights. *In-memory processing*, also called *streaming*, tackles this problem by using large-space RAM and flash memory, since the processing of data that's in memory is faster than the processing of data that's stored in disk. And, faster hardware — processors and chips — is in development.

WHAT'S THE FUTURE OF DATA-DRIVEN DECISION MAKING?

DDDM has come a long way in a short time and is likely to continue evolving at a breakneck pace. Particularly for tech companies, data acquisition is driving ever-faster innovation cycles as well as response times to market conditions. These trends raise interesting questions about the future of DDDM.

Tech companies have been leaders in pushing the limits of DDDM. In early 2018, Twitter CEO Jack Dorsey said the network had received questions about whether it was possible to measure the “health” of conversations on Twitter; the company began exploring data solutions that could make these judgments.

The impact of DDDM on the work experience and satisfaction of humans is receiving increased attention because of the larger social implications. A study by researchers at Carnegie Mellon University found that [people whose work assignments were driven by algorithms](#) — in this case, drivers for ride-share apps — experienced dissatisfaction when the assignments didn't make sense to them or made them feel they did not have enough control. The drivers felt that some algorithmically driven decisions “failed to accommodate their abilities, emotion, and motivation.”

The rise of automation in DDDM, a trend that is accelerating due to artificial intelligence, is also stirring debate. A good example of automated data-driven decision making is ride-sharing app Uber's dynamic pricing, which puts demand data into an algorithm to determine when premium fares will prevail.

Automation removes the possibility of human bias and human error, but also takes away the soft skills that humans provide. Automated price surges at Uber have occasionally gone into effect at inopportune times, such as after a terrorist attack in London in 2017. The company suspended the fare increase when it became aware of the incident, but consumers had already lashed out at the company on social media, accusing it of profiting from the attack.



As Ketan Kapoor, CEO and Co-Founder of human resources technology company Mettl, says, “You can certainly automate tasks that have clearly defined input parameters and no grey areas. If estimates or grey areas are involved anywhere in the process, human intervention is required.”

He continues, “But, with the advent of AI, as you keep feeding the system with complex-data points, systems in the future will ... deliver accurate inferences from the supplied data, irrespective of how complex the data is ... It will take time to shape up. However, the day is not far off when human intervention will be merely a redundant layer in the DDDM.”

Nathan of Derwen.ai recommends viewing machine-led and human-led decision making as a continuum: “Every business problem should now be viewed as something to be solved by teams of people plus machines. In some cases, there will be mostly people, while in others, there will be mostly machines. Still, you should always start from a baseline of managing teams of both.”

In human resources, analytics allows companies to examine things like headcount trends and worker demographics across units, departments, and organizations. By cross-referencing this data with desirable outcomes, such as productivity, analysts can look for relationships between demographics and performance.

Education is a growing focus for DDDM. Schools amass mountains of student performance data, using quarterly report cards and standardized testing. Some educators are analyzing this data for indications of which specific instructional methods (or teachers) have the greatest positive impact on learning. Then, these teachers are adjusting their teaching practices accordingly. These analyses can be simple to conduct if a school system is willing to invest in data expertise. Some helpful studies on DDDM in education include one by Michael J. Donhost and Vincent A. Anfara Jr. in [Middle School Journal](#) and another by Marlow Ediger in [College Student Journal](#).

With many schools providing students with tablets, which can function as data trackers, analysts have the potential to look at numerous variables, in addition to test scores. Educators can map personalized data and examine associations, such as whether the time spent on certain activities predicts performance improvements for low-achieving students or which instructional materials engage students the most, based on how many times they look away from the screen.

DATA-DRIVEN DECISION MAKING INCREASES DEMAND FOR DATA SCIENTISTS AND SPECIALISTS

As data-driven decision making takes hold, there is an increasing demand for people who have trained in data analysis tools and techniques.

Most major U.S. cities have a shortage of people with data science skills. In fact, as of late 2018, there was a national shortage of more than 150,000 people, according to LinkedIn. The shortage is especially acute in New York City and the San Francisco Bay Area. The U.S. Bureau of Labor Statistics predicts that by 2026, 11.5 million more jobs will be created in data science and analytics.

Website KDNuggets, which has an audience of data scientists, estimates that in 2018, there were already between 200,000 and 1 million people worldwide doing data science-related work. However, a 2015 poll of its members found that most expected expert-level predictive analytics and data science tasks to be automated by 2025. To remain in demand, data specialists should focus on cultivating skills that are hard to automate, like business understanding and storytelling.

According to InsideBigData.com, 88 percent of data scientists have master’s degrees, and 46 percent have

doctoral degrees. The talent shortage is exacerbated by a lack of university programs in data science, and that is prompting 63 percent of the companies that need these skills to provide in-house training.

Even if you aren't ready to embark on a degree program in data science, you can improve your DDDM expertise with a variety of courses and certificate programs. Check out some of the leading options below, including online, self-paced, free, and classroom alternatives.

Course Name

Platform

Taught by

Free/Paid

[Data-Driven Decision Making](#)

Coursera (online)

Alex Mannella, PwC

Paid (but financial aid available)

[Business Analytics for Data-Driven Decision Making](#)

EdX (online)

John W. Byers, Chris Dellarocas, Andrei Lapets, Boston University

Free (without certificate, paid for certificate)

[Data-Driven Decision Making](#)

Management Concepts (live classroom, virtual classroom)

-

Paid

[Data-Driven Decision Making](#)

Hyper Island In Person (international, in person)

-

Paid (with waivers available)

[Data Analytics Bootcamps](#)

Northeastern University (live classroom, online)

-
Course Name

Paid
Platform

Taught by
[Data Science: Data-Driven Decision Making](#)

Free/Paid
Learn@Forbes (online)

Frank Kane, Sundog Education

Paid

[Data-Driven Decision-Making](#)

Duke's Fuqua School of Business (in person)

Alexandre Belloni, Peng Sun, Saša Pekeč

Paid

[Data Driven Decision Making](#)

Udemy (online)

Douglas Clark

Paid

[Data-Driven Decision Making Certificate Program](#)

Cleveland State University (online)

-

Paid

[Business Analysis for Executives: Leveraging Data as a Strategic Asset](#)

NYU Stern School of Business (in person)

J.P. Eggers

Paid

[Data-Driven Decision Making](#)

Michigan Virtual

-

Paid

Course Name

[Data Science and Big Data Analytics: Making Data-Driven Decisions](#)

Platform

MIT (online)

Taught by

Devavrat Shah, Philippe Rigollet, et al.

Free/Paid

Paid

[Business Analytics](#)

Cornell (online)

Donna Haeger

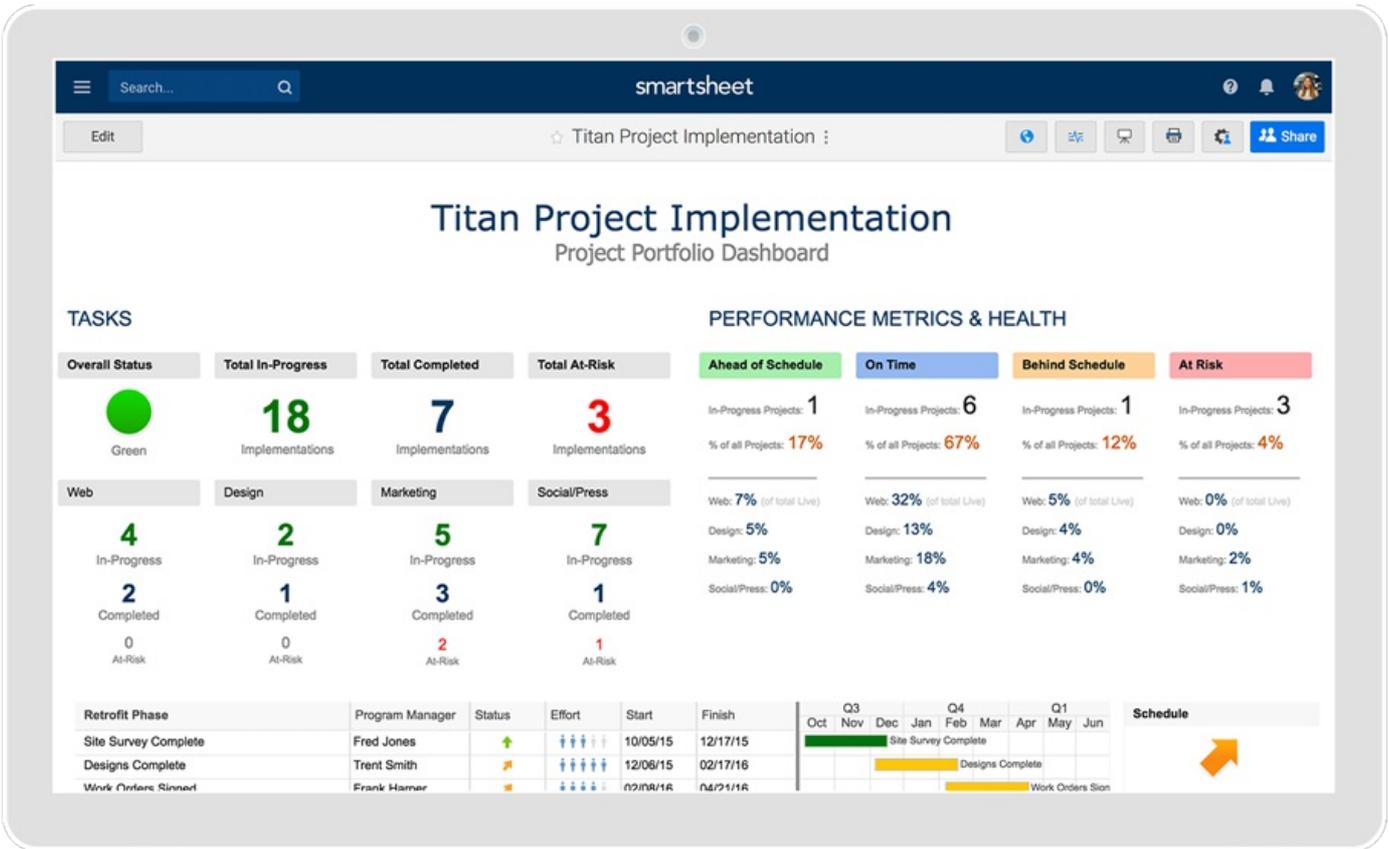
Paid

IMPROVE DATA-DRIVEN DECISION MAKING WITH SMARTSHEET FOR PROJECT MANAGEMENT

Top project managers rely on data-driven decision making to increase agility, detect new business opportunities, and respond to market changes more quickly — in essence, to remain competitive. But in order to achieve these goals, you need have the right tools.

Smartsheet is a work execution platform that enables enterprises and teams to get from idea to impact - fast. Top project management leaders rely on Smartsheet to help align the right people, resources, and schedules to get work done.

Use Smartsheet to create consistent project elements, increase speed, and improve collaboration with scalable options that fit individual work preferences. Hold yourself and your team accountable, improve visibility into team priorities, and ensure nothing slips through the cracks.



Try Smartsheet for Free

ADD NEW COMMENT

Your name

Subject

Comment *

Our [Privacy Policy](#) describes how we process your personal data.

Save



Smartsheet ▾

Product ▾

Solutions ▾

Support ▾

Legal ▾



©2019. All Rights Reserved Smartsheet Inc. Patents and Patents Pending.